Regulatory Elements in the HIV-1 5‘UTR that Modulate Gag Binding Specificity

Research Thesis

Presented in partial fulfillment of the requirements for graduation
with research distinction in Biochemistry in the undergraduate colleges of The Ohio State University

by

Joshua-Paolo C. Reyes

The Ohio State University

April 2018

Project Advisor: Dr. Karin Musier-Forsyth, Department of Chemistry and Biochemistry

TABLE OF CONTENTS

ABSTRACT

The 5'UTR of the HIV-1 genomic RNA (gRNA) contains a structured RNA element (termed Psi) that is specifically recognized through the interaction with the nucleocapsid (NC) domain of HIV-1 Gag, ensuring that two copies of gRNA are packaged into newly assembled virions. However, the mechanism by which Gag recognizes gRNA over other cellular RNAs and spliced RNAs is not well understood. Exposed guanosine residues within the loops and bulges of Psi have been shown to be high affinity NC binding sites, providing a selective advantage for gRNA. Furthermore, a recent study suggested that a negative regulatory element upstream of Psi reduces high-affinity Gag binding, and a positive downstream regulatory element counteracts the upstream element and restores high-affinity binding. These elements are proposed to form a long-range interaction that promotes packaging of only full-length gRNA and excludes packaging of spliced RNAs. Using a fluorescence anisotropy-based salt-titration binding assay, which measures the electrostatic and nonelectrostatic (i.e., specific) components of protein binding RNA, we have previously shown that Gag interacts with a 109 nt Psi RNA construct with high specificity and relatively few electrostatic interactions. High specificity of this Psi construct is maintained when guanosine residues in the bulges of SL2 are mutated, but guanosine mutations in the bulges of SL1 and regions downstream of SL3 greatly diminish this high specificity. When using a 356-nt RNA construct that now includes the upstream negative regulatory element in addition to Psi, we observed similar Gag binding specificity and electrostatic interactions to the 109 nt WT Psi construct. However, a 400 nt construct that additionally contains the positive element, increased specific binding and reduced the electrostatic interactions. Interestingly, Gag binding specificity remained high even in constructs

wherein the 44 nt of the downstream positive regulatory element were scrambled. Thus, while the presence of downstream sequences is important for high-specificity Gag binding, their identity is not. Furthermore, the secondary structures of spliced RNA constructs were probed via selective 2' hydroxyl acylation analyzed by primer extension (SHAPE) to study how they prevent specific packaging. In a construct composed of the first 400 nt of the Vpr mRNA, destabilization of the polyA hairpin was observed which has been shown to be critical for RNA packaging and viral replication. Meanwhile, the secondary structure of a construct containing the first 400 nt of the Tat mRNA displays a sequestered primer binding site (PBS) which has been shown to be necessary for $tRNA^{Lys3}$ annealing and initiation of reverse transcription. Altogether, these results allow us to better understand the mechanism by which genomic RNA packaging occurs.

INTRODUCTION

According to the World Health Organization, approximately 37 million people are currently infected by the Human Immunodeficiency Virus type 1 (HIV-1) worldwide, making it one of the most critical global health issues of today. A vaccine or cure for HIV-1 have yet to be developed, but with the advent of highly active antiretroviral therapy (HAART) the disease has been rendered into a manageable infection. HAART combines a variety of antiretroviral drugs that block HIV-1 replication. As a result of HAART, many treated patients are able to live healthy lives, symptom-free. However, the continuous use of antiretroviral medication may have other adverse effects and drives the risk of emergence of drug-resistant strains, thus forcing the continuous development of novel antiretroviral drugs. These drugs target many different stages in the HIV-1 retroviral lifecycle which include fusion, reverse transcription, integration, and maturation (Figure 1). Interestingly, there are no current antiretroviral drugs that target HIV-1 packaging and assembly, making it of significant interest in our lab and others.

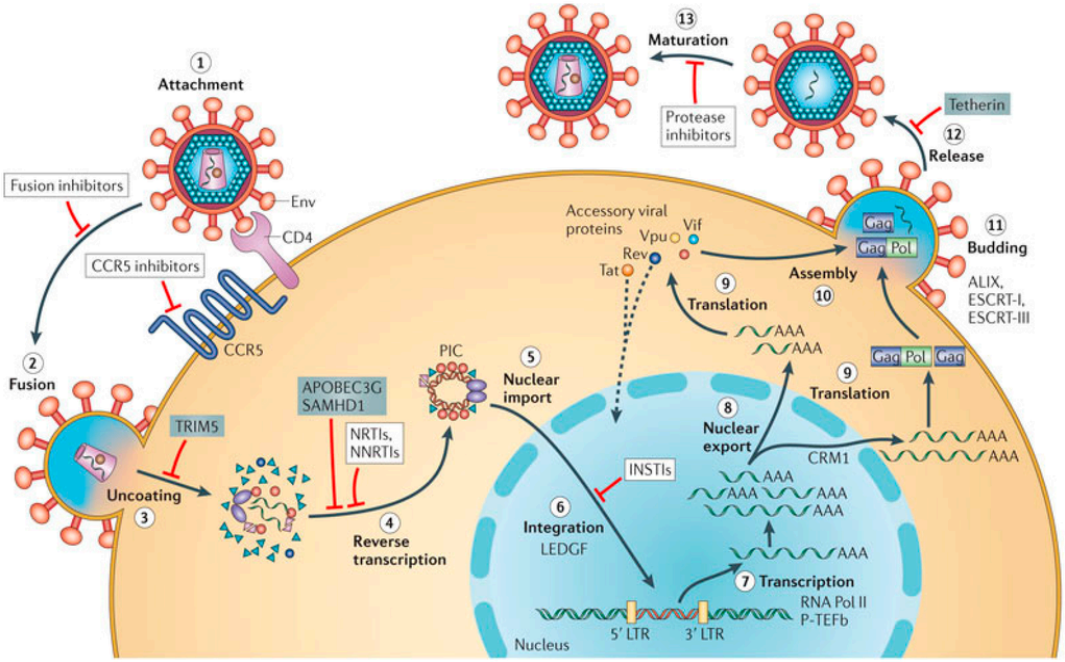


FIGURE 1: HIV-1 Lifecycle. Targets of current antiretroviral medications are highlighted in red (Musumeci, Riccardi and Montesarchio 2015).

A hallmark of retroviruses is their ability to reverse transcribe single-stranded genomic RNA (gRNA) into double-stranded DNA for subsequent integration into the host cell genome. Following nuclear transcription and export into the cytoplasm, highly selective packaging of the full-length, unsliced gRNA into newly assembled virus particles is achieved, despite the presence of spliced viral RNA and a vast excess of cellular RNA (Berkowitz, Fisher and Goff 1996, Kuzembayeva et al. 2014). HIV-1 packages only a single, dimeric copy of its gRNA (Rein et al. 2011, Moore and Hu 2009, Chen et al. 2009, Nikolaitchik et al. 2013). Specific interactions between the HIV-1 5' untranslated Region (5'UTR) and the Gag protein are critically involved in this specific packaging and assembly step of budding virions. The 5'UTR is the most highly conserved region in the viral genome and is also responsible for numerous aspects of viral life cycle regulation. It is comprised of the trans-activation response (TAR) stem loop, the poly-A hairpin, the primer binding site (PBS) domain, and the four stem loops (SL1-SL4) of the "Psi" (Ψ) packaging signal. The TAR stem loop is involved in viral RNA transcription upregulation (Bannwarth and Gatignol 2005), and the PBS domain contains the site where $tRNA^{Lys,3}$, the reverse transcription primer, anneals (Mak and Kleiman 1997, Goldschmidt et al. 2002). SL1 contains the genomic RNA (gRNA) dimerization initiation site (DIS), SL2 is the major 5' splice donor, SL3 is a high-affinity Gag binding site, and SL4 harbors the Gag gene start codon. The 5'UTR is also known to adopt multiple conformations allowing for preferential selection by HIV-1 Gag (Heng et al. 2012, Abbink and Berkhout 2003).

Experimental data suggest that the NC domain of Gag interacts with the gRNA in the 5'UTR at Psi. Specific HIV-1 gRNA packaging is achieved, in part, via the interaction between the nucleocapsid (NC) domain of the HIV-1 Gag polyprotein and Psi (Heng et al. 2012, Kutluay et al.

2014). In addition to the NC domain, HIV-1 Gag consists of matrix (MA), capsid (CA), two spacer peptides, and p6 (Figure 2). In the context of the polyprotein, MA targets the assembling virus to the plasma membrane (PM) (Chukkapalli and Ono 2011), CA is involved in Gag multimerization associated with immature viral particle formation (Datta and Rein 2009, Schur et al. 2015), NC functions as a nucleic acid chaperone in addition to its role in gRNA packaging (Levin et al. 2005, D'Souza and Summers 2005, Lu, Heng and Summers 2011, Feng et al. 1999, Jones et al. 2011, Cristofari and Darlix 2002, Rein, Henderson and Levin 1998), and the p6 domain recruits factors required for viral fission from the cell (Martin-Serrano and Neil 2011, Meng and Lever 2013). Importantly, the main structural domains of Gag provide functional redundancy to ensure proper viral assembly through Gag-nucleic acid, Gag-Gag, and Gag-membrane interactions (O'Carroll et al. 2012, O'Carroll et al. 2013).

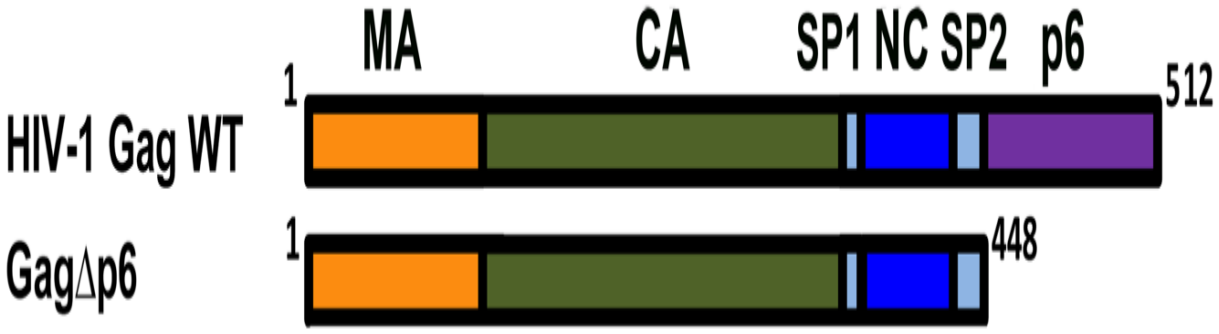


FIGURE 2. HIV-1 GagΔp6 construct used in this work with full-length wild-type (WT) Gag shown at the top of each set for comparison.

Furthermore, HIV-1 Gag has been shown to adopt multiple conformations in solution (Datta et al. 2011, Munro et al. 2014). One of the key functions of Gag is to ensure that the two strands of gRNA are properly incorporated into the new virions. However, the precise mechanism Gag employs to select two copies of gRNA over the vast excess of host and non-genomic viral RNA remains an open question. In this thesis, using *in vitro* biochemical techniques, I will

investigate the structural elements within the HIV-1 5'UTR that we hypothesize regulate specific Gag binding. This work will improve our understanding of HIV-1 gRNA packaging and may uncover new avenues for rational drug design.

CHAPTER 1: CRITICAL GUANOSINE RESIDUES IN HIV-1 PSI ARE REQUIRED FOR SPECIFIC RECOGNITION BY HIV-1 GAGΔP6

INTRODUCTION

Although the minimal element that is both necessary and sufficient to confer selective HIV-1 gRNA packaging is not yet known, a gRNA packaging competition assay has defined a "core encapsidation signal" composed of the U5 region and stem-loops 1-4 (SL1-SL4) of Psi (Heng et al. 2012). In the absence of Psi, efficient virion formation still occurs and cellular RNAs are incorporated into virions at levels corresponding to their relative amounts in the cytoplasm (Rulli et al. 2007). This indicates that Psi provides a "packaging advantage" to gRNA, and in the absence of this signal, random RNAs are passively incorporated into viral particles. Within the stem-loops that constitute Psi, exposed G residues in loops and bulges have been shown to be high-affinity NC binding sites based on a number of lines of evidence including selective 2'-hydroxyl analyzed by primer extension (SHAPE) footprinting studies of gRNA in virions (Wilkinson et al. 2008), *in vitro* RNA binding and footprinting assays (Keane et al. 2015, Abd El-Wahab et al. 2014), and crosslinking-immunoprecipitation-sequencing (CLIP-seq) techniques in HIV-1 infected cells (Kutluay et al. 2014).

Several factors in addition to Gag/NC binding affinity to Psi have recently been proposed to play a role in selective gRNA packaging. *In vitro* fluorescence-anisotropy (FA) salt-titration binding assays showed that HIV-1 Gag binding to RNA derived from the gRNA Psi region is characterized by highly specific non-electrostatic binding and an effective charge that is consistent with an NC-only binding mode (Webb et al. 2013). In contrast, binding to a non-Ψ RNA is characterized by less specific binding and a higher effective charge, consistent with nucleic acid

binding by both NC and MA domains. Thus, changes in Gag's binding conformation may contribute to the high degree of discrimination between Psi and non-Psi RNA (Webb et al. 2013).

As mentioned earlier, previous FA salt-titration binding assays provided insights into specific Psi RNA recognition by HIV-1 Gag (Webb et al. 2013). Two key parameters are obtained using this approach, $K_{d(1M)}$, which describes the non-electrostatic component of binding (i.e., the extrapolated dissociation constant at 1 M salt), and Z_{eff}, the effective charge of the protein–nucleic acid interaction, which is the number of Na^+ cations displaced upon nucleic acid binding (Rye-McCurdy, Rouzina and Musier-Forsyth 2015). We previously reported that HIV-1 Gag binds to Psi RNA with a dramatically reduced $K_{d(1M)}$ and lower Z_{eff} relative to a non-Psi RNA (TARpolyA). These results suggest that Gag interacts with RNA using different binding modes; both the NC and MA domains bind to RNA in the case of TARpolyA, whereas binding to Psi RNA preferentially involves the NC domain at the expense of MA interactions. Mutations altering the NC zinc finger motifs of Gag significantly reduce the non-electrostatic component of binding to Psi RNA and lead to an increase in Z_{eff}. Surprisingly, the specific binding of Gag to Psi also depended on the presence of the MA domain, as a Gag variant lacking MA was less able to discriminate Psi from non-Psi RNA and bound both RNAs with similar overall charge (Webb et al. 2013). We now investigated the RNA elements that contribute to specific recognition of Psi.

Materials and Methods

Protein Preparation

The plasmid encoding HIV-1 GagΔp6 was a gift from Dr. Alan Rein (HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute, Frederick, MD, USA). HIV-1 GagΔp6 was expressed and purified using established methods (Datta and Rein

2009). Protein concentrations were determined from the absorbance at 280 nm using the following molar extinction coefficient for HIV-1 GagΔp6: 63,090 $M^{-1} \cdot cm^{-1}$.

RNA Preparation

All RNA constructs were in vitro transcribed from linearized plasmids using T7 RNA polymerase and previously established methods (Milligan et al. 1987) (Figures 3A). The HIV-1 Ψ-WT RNA construct was derived from nt 229–338 and the TARpolyA RNA construct was derived from nt 1–104 of the HIV-1 NL4-3 genome (Adachi et al. 1986). Plasmids encoding both RNAs contain FokI restriction enzyme cut sites to generate the correct 3′ end of the RNA. All HIV-1 Ψ RNA variants were designed with 2 additional G residues at the 5′ end to promote efficient T7-mediated transcription. HIV-1 Ψ-Mut 1 (G241,242A), Ψ-Mut 2 (G274,275A), Ψ-Mut 3 (G292A, U293A, G294A), Ψ-Mut 4 (G311A), Ψ-Mut 5 (G319A, G321A), Ψ-Mut 6 (G329,330A), Ψ-Mut 7 (G235A), and Ψ-Mut 8 (G267,268A) were prepared from the HIV-1 Ψ-WT template plasmid using QuikChange (Agilent Technologies, Santa Clara, CA, USA) or site-directed ligase-independent mutagenesis (SLIM) (Chiu et al. 2004).

Fluorescent RNA Labeling

All RNAs were labeled with fluorescein-5-thiosemicarbazide (FTSC) at the 3′ end as described (Pagano et al. 2007). The concentration and labeling efficiency were determined by measuring the absorbance at 260 nm and 495 nm and using the following molar extinction coefficients: $\varepsilon_{495nm} = 8.5 \times 10^4\ M^{-1} \cdot cm^{-1}$ (fluorescein), $\varepsilon_{260nm} = 9.7 \times 10^5\ M^{-1} \cdot cm^{-1}$ (HIV-1 Ψ WT and HIV-1 Ψ mutants), and $\varepsilon_{260nm} = 9.3 \times 10^5\ M^{-1} \cdot cm^{-1}$ (HIV-1 TARpolyA).

Direct FA Binding Assays

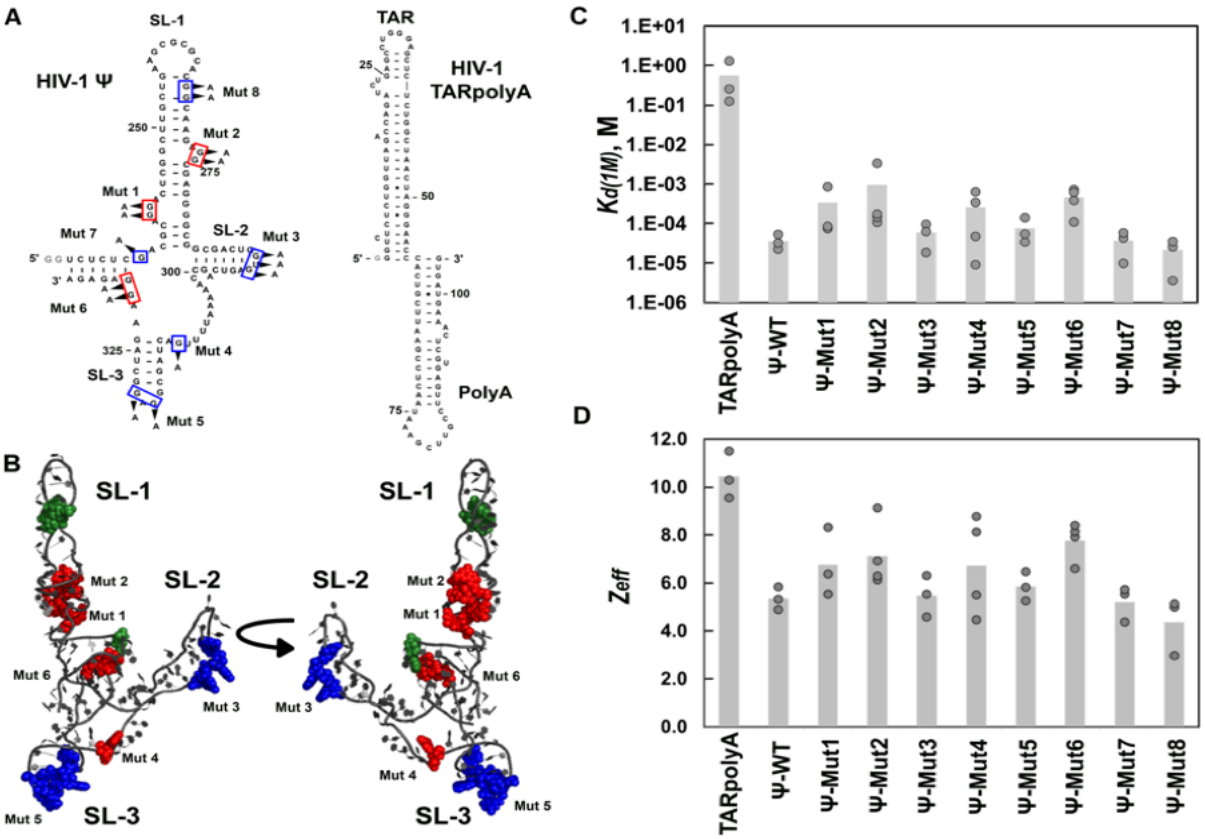
Direct FA binding assays were performed at 20 mM HEPES pH 7.5, 1 mM $MgCl_2$, 50 mM NaCl, and 20.5 nM RNA to obtain a preliminary assessment of binding affinity and to establish the concentration of each protein to use in the salt-titration studies as previously described (Webb et al. 2013, Rye-McCurdy et al. 2015).

FA-based Salt Titration Assays

FA salt-titration binding assays were carried out using established protocols (Webb et al. 2013, Rye-McCurdy et al. 2015). Assays were performed with 20.5 nM RNA in 20 mM HEPES or Tris-HCl, pH 7.5, 1 mM $MgCl_2$, and 50 mM monovalent ions with 300–400 nM HIV-1 GagΔp6. All fluorescence measurements were performed on a SpectraMax M5 plate reader (Molecular Devices, Sunnyvale, CA, USA).

Results and Discussion

In agreement with our previous work (Webb et al. 2013), HIV-1 GagΔp6 binding to WT Ψ RNA was characterized by an ~15,000-fold greater specificity ($K_{d(1M)}$ = 3.6 x 10-5 M) and ~6 fewer electrostatic interactions (Z_{eff} ~ 5) than binding to TARpolyA ($K_{d(1M)}$ = 5.6 x 10-1 M and Z_{eff} ~ 11) (Figure 3C and D). HIV-1 GagΔp6 exhibited only a minor loss in specificity upon mutation of residues in the loops of SL2 and SL3 of HIV-1 Ψ (Mut 3 and 5; $K_{d(1M)}$ ~ 6-8 x 10^{-5} M and relative specificity of 0.47-0.61 of Ψ-WT), and only a modest increase in the number of electrostatic charges mediating the interaction (Zeff ~ 5-6) (Figure 3C and D). Based on the partial overlap of the parameter values measured for these mutants with those for Ψ-WT, we conclude that these ~2-fold differences are not significantly different from WT. Similarly, although a point mutation in the single-stranded region between SL2 and SL3 (Mut 4) appeared to lead to a greater loss i



	HIV-1 GagΔp6		
RNA	$K_{d(1M)}$ (M)	Z_{eff}	Relative Specificity
TARpolyA	$(5.6 \pm 6) \times 10^{-1}$	10.5 ± 1.0	6.4×10^{-5}
Ψ-WT	$(3.6 \pm 2) \times 10^{-5}$	5.4 ± 0.5	1.0
Ψ -Mut1	$(3.4 \pm 5) \times 10^{-4}$	6.7 ± 1.4	0.11
Ψ -Mut2	$(9.5 \pm 16) \times 10^{-4}$	7.1 ± 1.4	0.038
Ψ -Mut3	$(5.9 \pm 4) \times 10^{-5}$	5.5 ± 0.9	0.61
Ψ -Mut4	$(2.6 \pm 3) \times 10^{-4}$	6.7 ± 2.1	0.14
Ψ -Mut5	$(7.6 \pm 6) \times 10^{-5}$	5.9 ± 0.6	0.47
Ψ -Mut6	$(4.6 \pm 3) \times 10^{-4}$	7.8 ± 0.8	0.078
Ψ –Mut7	$(3.7 \pm 2) \times 10^{-5}$	5.2 ± 0.7	0.97
Ψ –Mut8	$(2.1 \pm 2) \times 10^{-5}$	4.4 ± 1.2	1.7

FIGURE 3 HIV-1 RNA constructs used in this work and salt-titration binding results with HIV-1 GagΔp6. (A) Predicted secondary structures of HIV-1 Ψ (left) and HIV-1 TARpolyA (right) are shown. Eight mutant Ψ RNAs are indicated by the boxed nt and arrows (Mut1-Mut8). Additional 5' G nt not present in the HIV-1 genome were added to facilitate in vitro transcription and are shown in gray. (B) The location of Mut 1-8 mapped onto the all-atom model of HIV-1 Ψ previously determined from small angle x-ray scattering data and computational modeling. Colors in A and B indicate mutations that had no effect on specificity (blue), and significant effects (red) based on salt-titration binding assays. Bar graphs of Kd(1M) values (C) and Zeff values (D) determined from salt-titration assays with HIV-1 GagΔp6 and HIV-1 TARpolyA, Ψ RNA, and Ψ RNA mutants. Values of three or four trials performed in each case are shown with the height of the bar indicating the mean value. Specificity of the WT Ψ RNAs was set to 1.0 and the relative specificity of the non-Ψ or mutant Ψ RNAs was calculated as Kd(1M) (Ψ-WT)/ Kd(1M) (Ψ-variant).

specificity ($K_{d(1M)}$ ~ 3 x 10^{-4} M; relative specificity of 0.14), based on the variability in the data generated using this mutant, we conclude that it was not significantly different from Ψ-WT. In contrast, a significant loss in specificity was observed upon mutation of residues in the bulge regions of SL1 (Mut 1 and 2), as well as in the region downstream of SL3 (Mut 6) (relative specificity of 0.04-0.11). The final two HIV-1 Ψ mutants tested (Mut 7 and 8) bound HIV-1 GagΔp6 with a relative specificity of 0.97-1.7 and interacted with a similar number of electrostatic contacts (Z_{eff} ~ 4-5) as Ψ-WT (Figure 3C and D). While we observed different effects of the HIV-1 Ψ mutants on Gag binding specificity, no single Ψ mutant diminished specificity to the level of non-Ψ RNA.

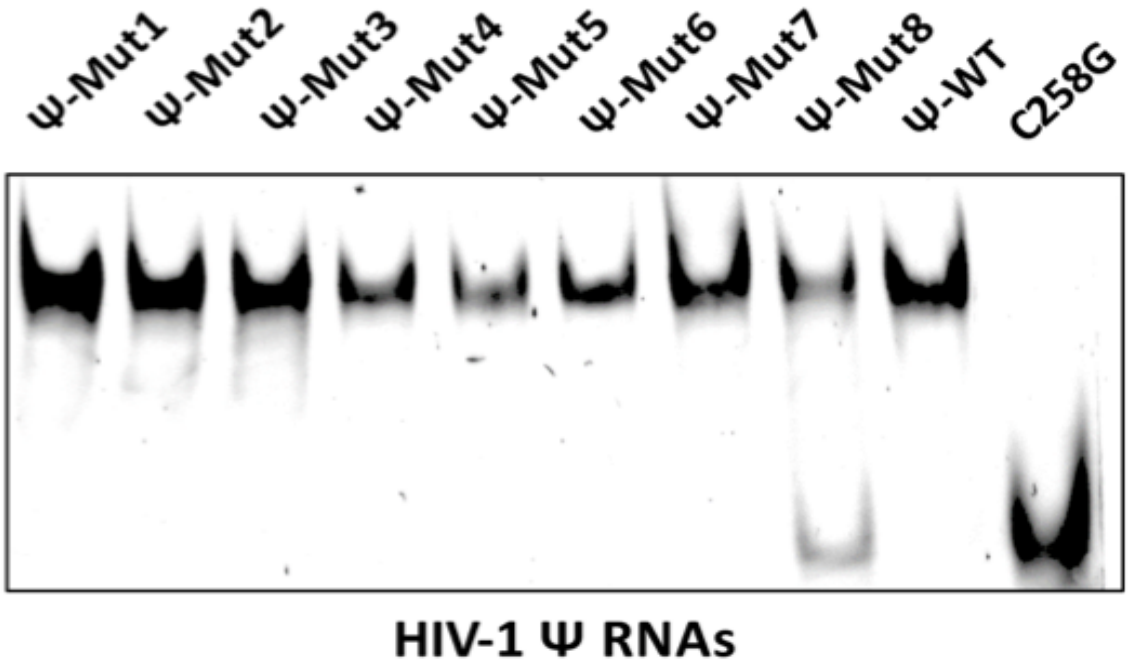


FIGURE 4 Native-PAGE of fluorescently-labeled WT and mutant HIV-1 Ψ RNAs. RNAs were labeled as described in the main text and folded as follows: 1-5 µM RNA in 50 mM HEPES, pH 7.5, was heated at 80°C for 2 min, then 60°C for 2 min, and then $MgCl_2$ was added to a final concentration of 10 mM, followed by incubation on ice for a minimum of 30 min. Folded RNAs were run on 8% native-polyacrylamide gels (19:1 acrylamide: bis-acrylamide) also containing 1 mM $MgCl_2$ and TB (18 mM Tris, 9 mM borate, pH 8) at 4°C in TB running buffer. The final concentration HIV-1 RNA run on the gels shown was 20.5 nM. Gels were visualized using the fluorescence mode on an AlphaImager EP (AlphaInnotech).

To determine whether the Ψ mutations disrupted the global RNA fold, all HIV-1 Ψ RNA variants were subjected to native-PAGE analysis. The majority of RNAs migrated as a single band corresponding to the RNA dimer, and there were no changes in the global conformation of the mutants relative to Ψ-WT (Figure 4). The only exception was Mut 8, which was a mix of dimer (major band) and monomer (minor band).

Previous salt-titration binding assays investigated a HIV-1 Ψ variant (Ψ-12M) wherein 12 single-stranded G bases in loops and bulges (clustered in 6 regions) were simultaneously mutated. These regions were proposed to be high-affinity NC binding sites based on a SHAPE footprinting study wherein the HIV-1 gRNA was probed in virio in the presence and absence of bound NC (Wilkinson et al. 2008). These mutations were performed in a slightly different Ψ RNA construct than the one used in the present work. Nevertheless, we found that HIV-1 Gag binding to the Ψ-12M RNA was characterized by a ~25-fold weaker non-electrostatic binding component relative to WT Ψ RNA (Webb et al. 2013). We now report the results of individual mutations at each of these 6 regions (1-3 mutations per region, see Figure 3A and 3B). Mutation of two G-rich bulges in SL1 (Ψ-Mut 1 and 2) and single-stranded G residues downstream of SL3 (Ψ-Mut 6) displayed the largest loss in specificity (Figure 3), with the mutation in the upper single-stranded bulge of SL1 (Mut2) having a similar effect as previously reported for Psi-12M (~25-fold decrease in specificity) (Webb et al. 2013). Interestingly, this bulge together with the lower bulge in SL1 (mutation of which reduces specificity ~10-fold in our assay) were found to be the two most statistically significant gRNA NC-interaction sites in the SHAPE study (Wilkinson et al. 2008). Our results are also consistent with a recent report identifying the upper bulge in SL1 (i.e., Ψ-Mut 2) as a critical site for full-length HIV-1 Gag interaction in vitro (Abd El-Wahab et al. 2014). Mutations

at 5 other locations in Ψ (Mut 3, 4, 5, 7, and 8) resulted in Gag binding specificity that was not distinguishable from Ψ-WT. Consistent with our results for Ψ-Mut 3, previous gRNA packaging studies found SL2 to be less important than the other elements of Ψ in directing gRNA incorporation into virions. (Kutluay et al. 2014, Lu et al. 2011, Keane et al. 2015, Abd El-Wahab et al. 2014, McBride and Panganiban 1996, McBride and Panganiban 1997). The Ψ-Mut 5 results were more surprising (~2-fold decrease in selectivity), as most studies agree on the importance of SL3 in packaging (Lu et al. 2011, McBride and Panganiban 1997). Importantly, mutation of a single-stranded G residue within Ψ that had not previously been implicated as a site of NC interaction (Ψ-Mut 7) did not negatively affect Gag binding specificity. Ψ-Mut 8 was the only HIV-1 RNA mutant tested to modestly disrupt dimerization activity (Figure 3), but consistent with our early report (Webb et al. 2013), this does not impact Gag binding specificity of these ~100-nt constructs. While a recent mutational interference mapping study examining the RNA residues critical for HIV-1 Gag binding to Ψ (Smyth et al. 2015) did identify the region corresponding to Ψ-Mut 8 as important, other studies are more consistent with a lack of direct Gag/NC binding in this region (Kutluay et al. 2014, Wilkinson et al. 2008).

In summary, mutational analysis of guanosine residues within Psi tested using salt-titration assays have indicated regions critical for Gag recognition. G residues in the bulges of SL1 and the single-stranded regions downstream of SL3 were found to be integral for Gag recognition, consistent with most studies. We mapped these three most important sets of residues for promoting specific Gag binding onto a tertiary structural model of HIV-1 Ψ RNA generated by small-angle X-ray scattering (SAXS) (Figure 3B) (Jones et al. 2014). Interestingly, these mutations cluster together in the central part of the RNA. Since no single-site mutant

reduces Gag binding specificity to the level of TARpolyA, we propose that the clustering of several Gag interaction sites in close proximity is an important determinant of specific binding.

ACKNOWLEDGEMENTS

I would like to thank Erik Olson for his patience and mentorship in this project as it was my first exposure to scientific research. I would also like to thank Brian Thompson for the generation of the Ψ-mutants and the analysis of preliminary salt titration data. Lastly, I would like to thank Shuohui Liu for his help preparing HIV-1 GagΔp6 and his help in testing the Ψ-mutants.

CHAPTER 2: HIV-1 5'UTR REGULATORY ELEMENTS MODULATE GAGΔP6 BINDING SPECIFICITY

INTRODUCTION

Gag plays a crucial role within the HIV-1 lifecycle by facilitating the packaging of two copies of gRNA over other vsRNAs and cellular RNAs. Both genomic and virally spliced RNAs (vsRNAs) harbor the critical 5'UTR elements TAR, polyA, PBS, and SL1, differing in their sequences downstream of the major 5' splice donor site within the hairpin loop of SL2. SL1 has been shown to be the main packaging determinant of gRNA, suggesting that vsRNAs avoid specific packaging through elements within their unique 3' sequences downstream of the major 5' splice site (Houzet et al. 2007). Further studies have hypothesized that a double regulatory mechanism within the HIV-1 5'UTR allow for specific selection of gRNA wherein a negative regulatory element that encompasses the TAR, polyA, and PBS stem loops, upstream of Psi, reduce high-affinity Gag binding, while a positive downstream regulatory element, only present in gRNA, counteracts the upstream element and restores high-affinity Gag binding (Abd El-Wahab et al. 2014) (Figure 5).

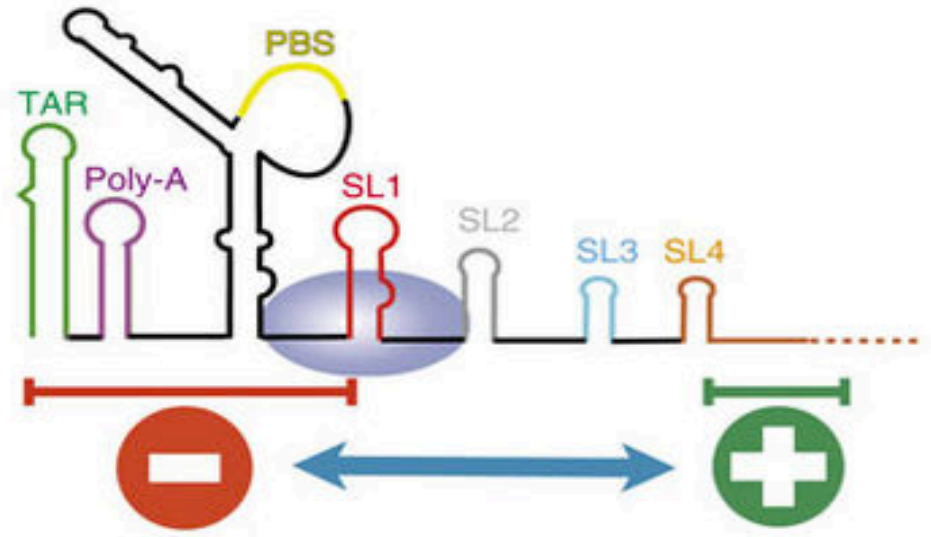


FIGURE 5. Schematic of the 5'UTR double regulatory system flanking the four stem loops of Psi. Upstream of SL1 is the negative regulatory element, underlined with a red minus sign, and downstream of SL4 is the positive regulatory element, underlined with a green plus sign. Figure from (Abd El-Wahab et al. 2014).

Although this study made important findings regarding the elements of gRNA that are required for packaging, it did not directly address the mechanism of how these RNAs affect Gag in a manner that would facilitate their packaging. Gag has been shown to adopt two distinct conformations when binding to Psi RNA and non-Psi RNA using salt titration analysis (Webb et al. 2013). Gag was shown to interact with Psi RNA in a linear conformation using primarily the NC domain while Gag interacted with non-Psi RNA in a bent conformation using both the NC and MA domains (Webb et al. 2013) (Figure 6).

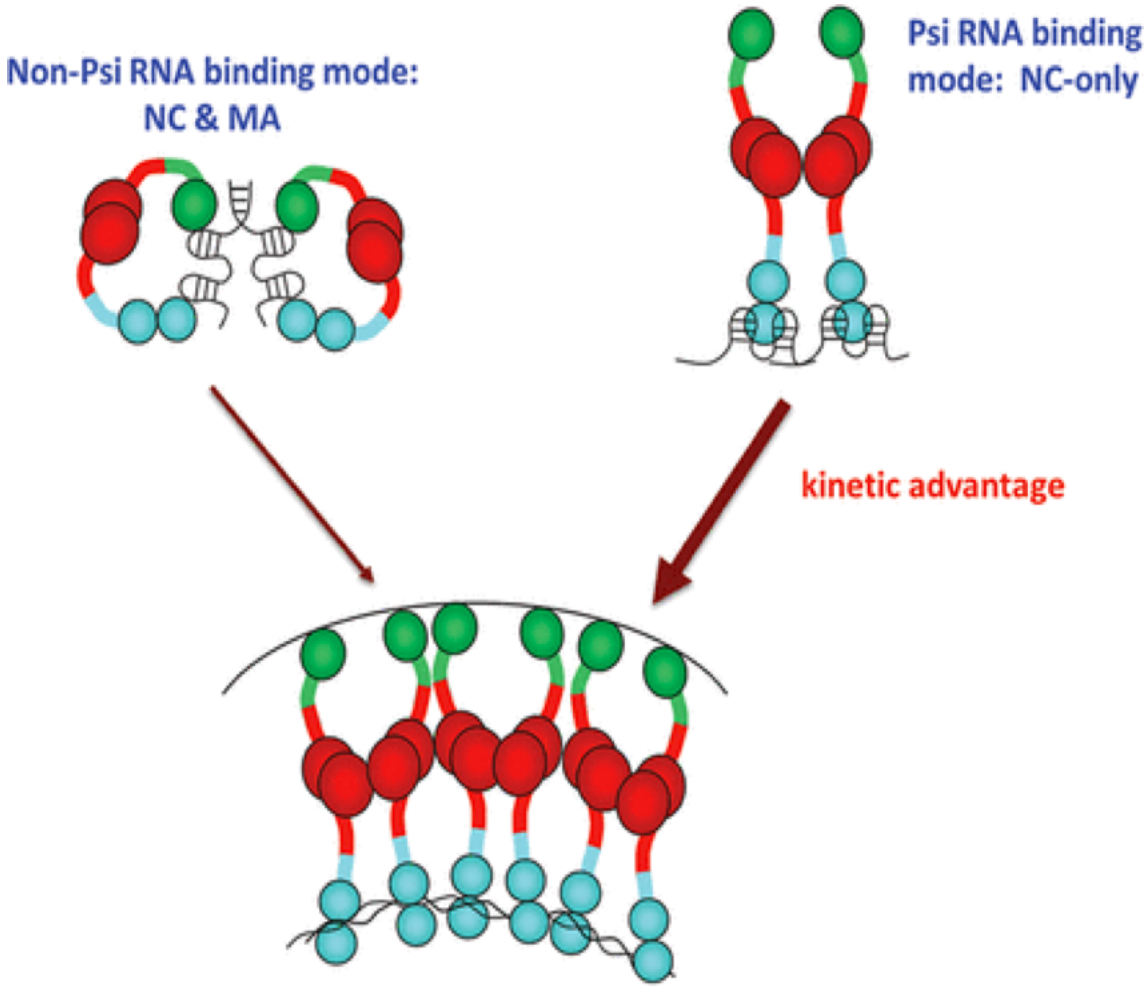


FIGURE 6. Model of Gag conformation upon binding to non-Psi and Psi RNAs. Gag binding to non-Psi RNAs is characterized in a NC and MA binding mode while Gag binds to Psi RNAs in an NC-only mode allowing it a kinetic advantage upon binding MA to the plasma membrane. Figure from (Webb et al. 2013).

Using the salt titration assay, we will test the hypothesis that Gag binds to RNAs containing the putative negative regulatory element and Psi stem loops in the bent conformation. In contrast, when binding to RNAs that contain the negative regulatory element, Psi, and the positive regulatory element, we hypothesize that Gag will to bind in a straight conformation similar to binding Psi alone.

Materials and Methods

RNA preparation

Nine RNA constructs containing different lengths of the HIV-1 NL4-3 5'UTR were studied using direct FA binding assays and FA salt titration assays (Figure 7). For preparation Ψ-WT and TARpolyA RNA constructs, see "RNA preparation" in Chapter 1. The 228mer, 356mer, and 400mer RNA constructs were derived from the first 228, 356, and 400 nucleotides (nt) of the HIV-1 genomic RNA, respectively. The Ψ+ construct contains nt 229–400 and the Omega (Ω) construct contains nt 356-400. Mutants of the 400mer construct were also used in this study wherein the last 44 nt of the RNA sequence were scrambled using a scramble function on an online Text Mechanic™ tool. This was performed twice, generating two mutants (400mer mut1 & mut 2). All RNAs were generated from plasmids encoding the T7 promoter sequence, the necessary RNA sequence, FokI restriction sites, and the Ampicillin resistance gene. Mutagenesis of plasmid constructs was performed using SLIM (Chiu et al. 2004).

Protein preparation

A functional mutant of the Gag polyprotein with a truncated C-terminal p6 domain, GagΔp6, was the primary protein used in this study and its more detailed preparation can be found in "Protein preparation" of Chapter 1.

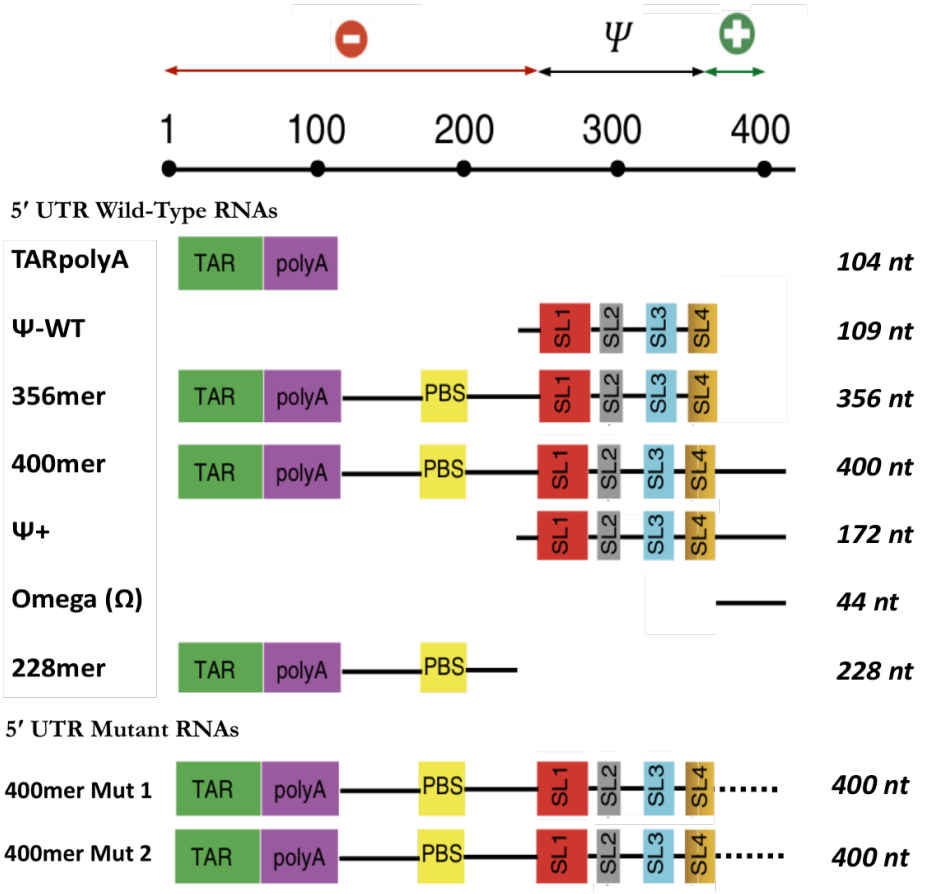


FIGURE 7. 5'UTR RNA Constructs used in this study. Dashed lines indicate scramble of the 44-nts in the putative positive regulatory element. Figure modified from (Abd El-Wahab et al. 2014).

FLUORESCENT RNA LABELING

All RNAs in this study were fluorescently labeled with FTSC at the 3' position as described in Chapter 1: "Fluorescent RNA Labeling". The concentration and labeling efficiency were determined by measuring the absorbance at 260 nm and 495 nm and using the following molar extinction coefficients: $\varepsilon_{495nm} = 8.5 \times 10^4\ M^{-1}{\cdot}cm^{-1}$ (fluorescein), $\varepsilon_{260nm} = 9.7 \times 10^5\ M^{-1}{\cdot}cm^{-1}$ (HIV-1 Ψ WT), $\varepsilon_{260nm} = 9.3 \times 10^5\ M^{-1}{\cdot}cm^{-1}$ (HIV-1 TARpolyA), $\varepsilon_{260nm} = 3.3 \times 10^6\ M^{-1}{\cdot}cm^{-1}$ (HIV-1 356mer), $\varepsilon_{260nm} = 3.7 \times 10^6\ M^{-1}{\cdot}cm^{-1}$ (HIV-1 400mer WT & 400mer mutants), $\varepsilon_{260nm} = 1.6 \times 10^6\ M^{-1}{\cdot}cm^{-1}$ (HIV-1 Ψ+), $\varepsilon_{260nm} = 3.7\times 10^5\ M^{-1}{\cdot}cm^{-1}$ (HIV-1 Ω), and $\varepsilon_{260nm} = 2.1 \times 10^6\ M^{-1}{\cdot}cm^{-1}$ (HIV-1 228mer).

DIRECT FA BINDING ASSAYS

Direct FA binding assays were first performed under the same conditions as described in "Direct FA binding assays" in Chapter 1 using all RNA constructs in Figure 7 with the exception of Ψ-WT and TARpoly A for which optimal protein concentrations have already determined (Webb et al. 2013, Rye-McCurdy et al. 2016). Optimal protein concentrations for each RNA were determined from direct FA binding assays by selecting near-saturating protein concentrations to minimize nucleic acid aggregation (Rye-McCurdy et al. 2015, Webb et al. 2013). Optimal GagΔp6 concentrations used for salt titration assays are as follows: 400 nM (HIV-1 Ψ-WT, Ψ+, TARpolyA), 600 nM (HIV-1 228mer), 1000 nM (HIV-1 Ω, 356mer, 400mer WT, & 400mer mutants). Direct FA binding assays were not directly performed on the 400mer mutants, as we assumed the constructs would have similar affinity to GagΔp6 as the 400mer WT.

FA-BASED SALT TITRATION ASSAYS

FA salt-titration binding assays were carried out using established protocols (Webb et al. 2013, Rye-McCurdy et al. 2015). Assays were performed with 20.5 nM RNA in 20 mM HEPES, pH 7.5, 2 mM Tris-HCl, 1 mM β-mercaptoethanol (βME), 1 mM $MgCl_2$, and 50 mM NaCl with 400–1000 nM HIV-1 GagΔp6. All fluorescence measurements were performed on a SpectraMax M5 plate reader (Molecular Devices, Sunnyvale, CA, USA).

RESULTS AND DISCUSSION

In this study, we used Ψ-WT and TARpolyA as our controls for specific binding and non-specific binding to GagΔp6, respectively. Relative to previously published data (Webb et al. 2013, Rye-McCurdy et al. 2016), the $K_{d(1M)}$ values measured here are approximately 10-fold less specific for both Ψ-WT and TARpolyA. Moreover, Ψ-WT bound with two more electrostatic contacts (Z_{eff}

~7) relative to previous reports (Webb et al. 2013, Rye-McCurdy et al. 2016). Nevertheless, The difference in specificity between Ψ-WT and TARpolyA in this study was ~10^4, which is similar to previous results (Rye-McCurdy et al. 2016).

GagΔp6 binding to constructs that lack Ψ but contain the negative element alone (228mer construct) results in $K_{d(1M)}$ and Z_{eff} values that are increased compared to the 356mer construct, but overall binding specificity is not as diminished as for the TARpolyA construct. Strikingly, when testing the positive element alone (Ω construct), we observe the highest binding specificity of all of the RNA constructs in this study, as well as the lowest Z_{eff}. The latter can be attributed to its short length at 44 nt. The secondary structure of Ω, predicted using an RNA secondary structure prediction program, Mfold (Zuker 2003), is shown in Figure 8. Based on this structure, the large number of single-stranded G residues, which are high-affinity NC binding sites, contributes greatly to the high binding specificity observed in this construct.

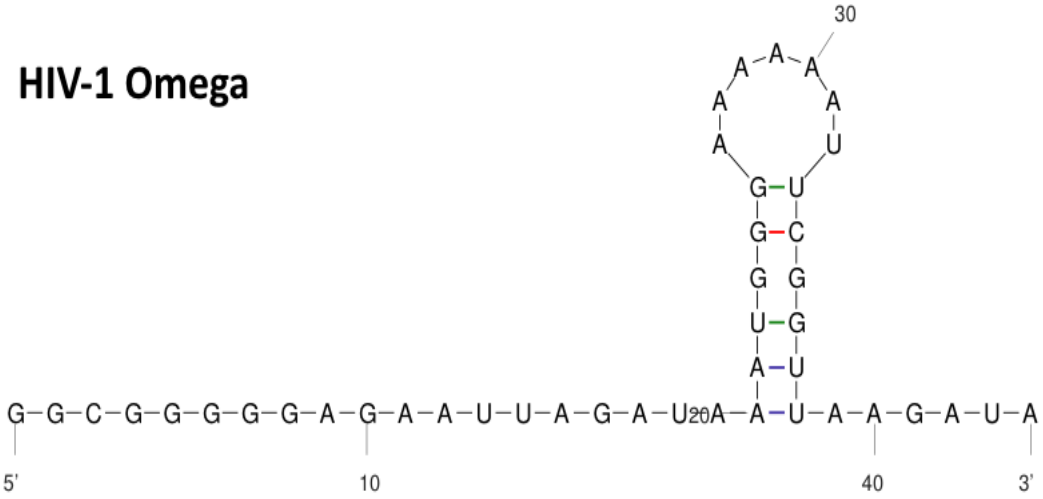
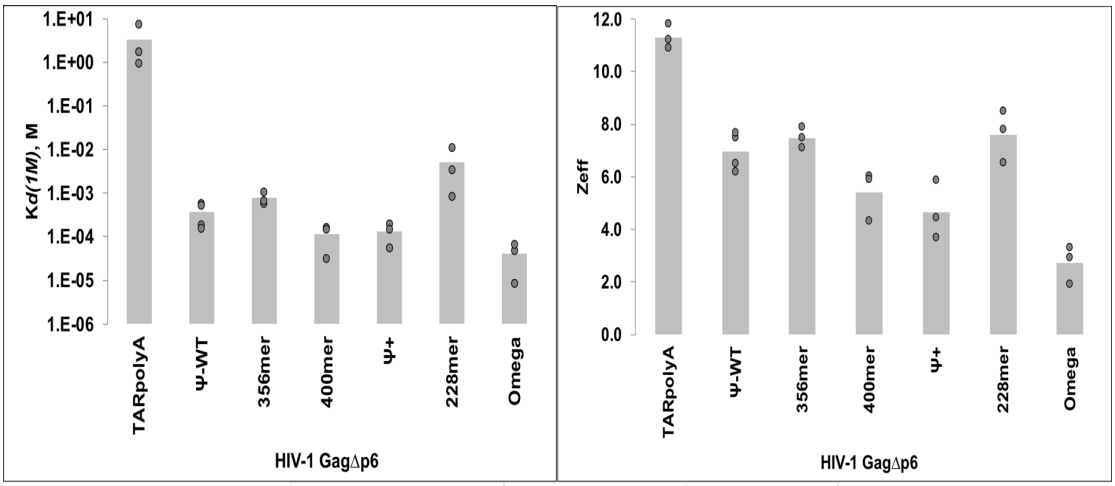


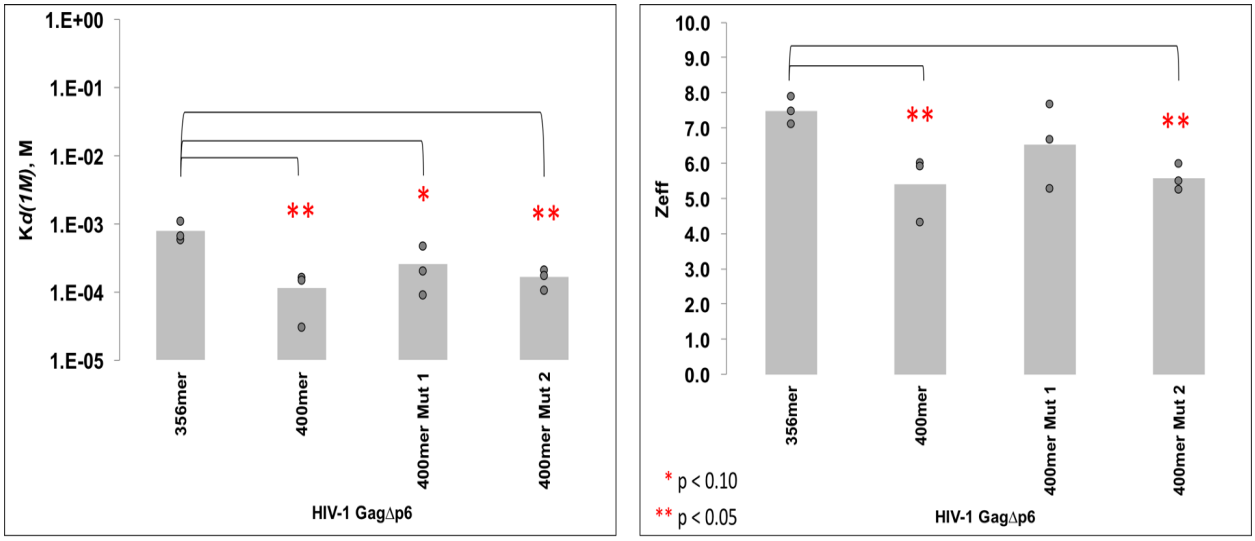
FIGURE 8. Secondary structure prediction of HIV-1 Omega predicted by Mfold (ΔG = -1.90 kcal/mol). RNA folding predicts 9 single-stranded G residues, which are high-affinity NC binding sites.

The 356mer, which consists of the putative negative regulatory element and SL1-SL4 of Ψ, results in only a modest increase in both the $K_{d(1M)}$ and Z_{eff} values when compared to the Ψ-WT construct. The overall decrease in specificity in the 356mer is approximately 2-fold. When testing the 400mer WT construct, which consists of Ψ and both putative, flanking regulatory elements, we observe a decrease in both $K_{d(1M)}$ and Z_{eff} values is observed relative to both Ψ-WT and the 356mer, indicating more specific recognition by GagΔp6 in the 400mer WT (Figure 9). Very similar parameters were measured for the Ψ+ construct, which contains SL1-SL4 of Ψ and the putative positive regulatory element. Thus, the addition of the positive element to both the 356mer and Ψ-WT increases GagΔp6 binding specificity both in the presence and absence of the negative element.



GagΔp6		
RNA Variant (20.5 nM)	**$K_{d(1M)}$**	**Z_{eff}**
TARpolyA	3.34 (± 3.50) x 10^{0}	11 (± 0.5)
Ψ-WT	3.64 (± 2.22) x 10^{-4}	7.0 (± 0.7)
356mer	7.77 (± 2.66) x 10^{-4}	7.5 (± 0.4)
400mer	1.14 (± 0.73) x 10^{-4}	5.4 (± 1.0)
Ψ+	1.34 (± 0.74) x 10^{-4}	4.7 (± 1.1)
228mer	5.15 (± 5.32) x 10^{-3}	7.6 (± 1.0)
Omega	4.14 (± 3.01) x 10^{-5}	2.7 (± 0.7)

FIGURE 9. Salt Titration Results of 5'UTR Wild Type RNAs. Bar graphs indicate Kd(1M) values and Zeff values determined from salt titration assays with HIV-1 GagΔp6. Average values and standard deviations of triplicated experiments are shown in the table below.



GagΔp6		
RNA Variant (20.5 nM)	**$K_{d(1M)}$**	**Z_{eff}**
356mer	7.77 (± 2.66) x 10^{-4}	7.5 (± 0.4)
400mer	1.14 (± 0.73) x 10^{-4}	5.4 (± 1.0)
400mer Mut 1	2.56 (± 1.96) x 10^{-4}	6.5 (± 1.2)
400mer Mut 2	1.62 (± 0.53) x 10^{-4}	5.6 (± 0.4)

FIGURE 10. Salt Titration Results of the HIV-1 356mer, 400mer, and 400mer mutant RNAs. Bar graphs indicate Kd(1M) values and Zeff values determined from salt titration assays with HIV-1 GagΔp6. P-values obtained from T-tests between the 356mer and the 400mer mutants are shown by brackets followed by a "*" indicating $p < 0.10$ and "**" indicating $p < 0.05$. Average values and standard deviations of triplicated experiments are shown in the table below.

To further investigate the putative positive regulatory element and to establish whether its effect in conferring greater Gag binding specificity are sequence dependent, 400mer mutants were generated containing different scrambled 44-nt sequences that comprise the positive element. Salt titration analysis (Figure 10) revealed a two-fold decrease in the binding specificity between 400mer WT and 400mer mut 1 but no change between 400mer WT and 400mer mut 2. T-tests were used to compare the $K_{d(1M)}$ and Z_{eff} values of the 400mer WT and the mutants and no statistically significant differences were found ($p \not< 0.05$) for either mutant. However, when t-tests were performed to compare the 356mer and the 400mer mutants, statistically significant differences were found in both the $K_{d(1M)}$ and Z_{eff} values for 400mer mut 2 ($p < 0.05$). Overall,

these data suggest that the putative positive regulatory element in the HIV-1 5'UTR functions independent of the sequence of nucleotides.

Taken together, the data reported here do not provide evidence for interactions between the reported positive and negative regulatory elements in the HIV-1 5'UTR. Our data suggest that the positive regulatory element is able to increase GagΔp6 binding specificity, independent of the negative regulatory element or its sequence. Given that sequences downstream of Psi, independent of sequence, are sufficient to facilitate an increase in Gag binding specificity, these results open up new questions into the mechanisms underlying genomic RNA packaging. In particular, the question of what elements within in the virally spliced RNAs disrupt their specific packaging into virions will be explored in the following chapter.

ACKNOWLEDGEMENTS

I would like to thank Erik Olson for his guidance on this project and his trust in my ability to work independently under his supervision. I would also like to thank Will Cantara and Jonathan Kitzrow for their general insight and suggestions on the project.

CHAPTER 3: SECONDARY STRUCTURE OF SPLICED HIV-1 TRANSCRIPTS REVEALS POTENTIAL MECHANISMS TO PREVENT PACKAGING.

INTRODUCTION

The HIV-1 full-length RNA transcript is approximately 9-kb and encodes for 15-18 viral proteins from both full-length and alternatively-spliced transcripts. Full-length mRNA transcripts are selectively packaged into budding virions and are also translated to produce the Gag and Gag-Pol proteins. Approximately half of the viral RNAs exported out of the nucleus are full-length transcripts (Stoltzfus 2009). Full length mRNA transcripts can be singly spliced, producing mRNA transcripts for the Vif, Vpr, Vpu, and Env proteins, or multiply spliced, producing the Rev, Tat, and Nef proteins. The proteins produced from multiply-spliced mRNAs are further characterized by their 1.8-kb transcript and their translation in the early phase of HIV-1 infection (Stoltzfus 2009). In contrast, the proteins produced from the singly-spliced transcripts are further characterized by their 4-kb transcript and their translation in the late phase of infection (Stoltzfus 2009). The HIV-1 5'UTR plays a significant role in the production of spliced viral mRNAs as it houses the major 5' splice site in SL2 used by both singly- and multiply- spliced transcripts. Thus, vsRNAs all contain identical 5' sequences that include TAR, polyA, PBS, and SL1 and differ in sequences past the major 5' splice site. These vsRNAs avoid specific packaging into budding virions by Gag despite harboring sequences like SL1, which have been shown to be necessary for efficient packaging (Aldovini and Young 1990, Houzet et al. 2007, Abd El-Wahab et al. 2014). We hypothesize that structural differences between vsRNAs and full-length gRNA in the 5'UTR may serve to block packaging of the former. To probe this idea, the secondary structure of the first 400 nt of the gRNA, as well as the first 400 nt of the Tat and Vpr mRNA transcripts, was determined using

selective 2′ hydroxyl acylation analyzed by primer extension (SHAPE).

Materials and Methods

RNA Preparation

Four RNA constructs were used in this study: HIV-1 400mer wt, 400mer ΔDIS, Tat 400mer, and Vpr 400mer (Figure 11). Preparation of the 400mer wt construct can be found in "RNA Preparation" of Chapter 2. In the 400mer ΔDIS construct, the dimerization initiation signal (GCGCG) sequence in the loop of SL1 was replaced with a stable GAGA tetraloop to prevent RNA dimerization. Overlap extension PCR was performed on the 400mer wt plasmid to generate the mutation (Bryksin and Matsumura 2010). The Tat and Vpr 400mer constructs were designed from similar constructs found in (Bernacchi et al. 2017). Vpr 400mer is comprised of nt 1-290 of the HIV-1 NL4-3 genome and the first 110 nt of the Vpr exon. Tat 400mer is also comprised of nt 1-290 of the HIV-1 NL4-3 genome but has the first 110 nt of the Tat exon. Plasmids containing the T7 promoter, RNA sequence of interest, FokI restriction cut sites, and Ampicillin resistance gene were designed and purchased from Integrated DNA Technologies, Inc. Subsequent RNA preparation steps can be found in "RNA preparation" Chapter 1.

Fluorescent RNA Labeling

All RNAs in this study were fluorescently labeled with FTSC at the 3′ position as described in Chapter 1: "Fluorescent RNA Labeling". The concentration and labeling efficiency were determined by measuring the absorbance at 260 nm and 495 nm and using the following molar extinction coefficients: $\varepsilon_{495nm} = 8.5 \times 10^4\ M^{-1}\cdot cm^{-1}$ (fluorescein), $\varepsilon_{260nm} = 3.69 \times 10^6\ M^{-1}\cdot cm^{-1}$ (HIV-1 400mer WT, Vpr 400mer, and Tat 400mer), and $\varepsilon_{260nm} = 3.66 \times 10^6\ M^{-1}\cdot cm^{-1}$ (HIV-1 400mer ΔDIS).

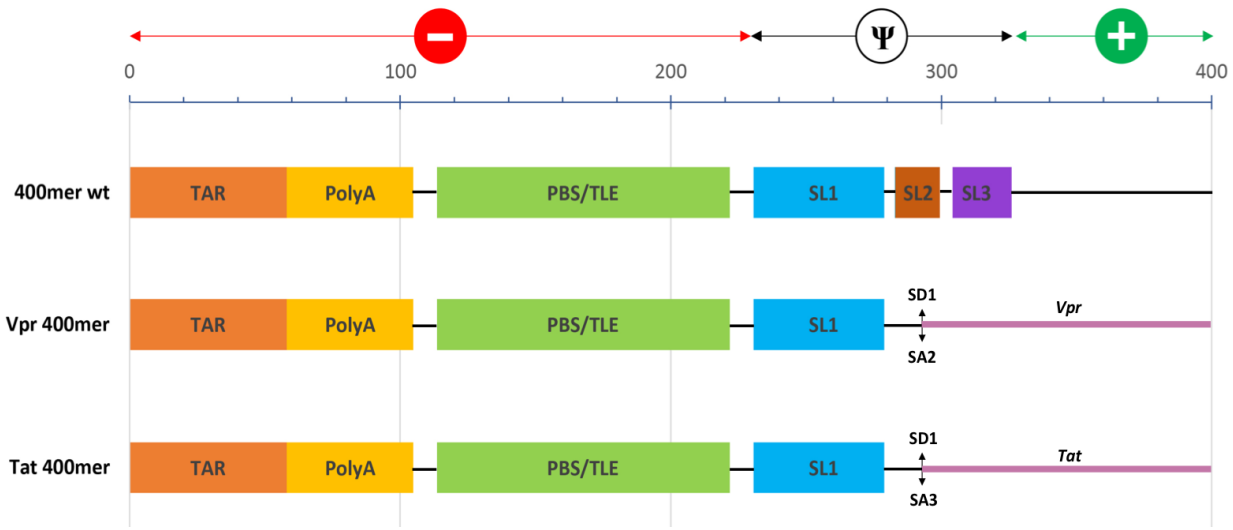


FIGURE 11. RNA constructs used in this study. HIV-1 400mer ΔDIS is not shown in this figure. The purple lines downstream of the major 5' splice site (SD1) indicate sequences from splice acceptor sites (SA2 & SA3) of Vpr and Tat.

SELECTIVE 2' HYDROXYL ACYLATION ANALYZED BY PRIMER EXTENSION (SHAPE)

Unlabeled 400mer WT, Vpr 400mer, and Tat 400mer RNAs were probed using SHAPE (Wilkinson, Merino and Weeks 2006). RNAs were first folded as follows: 0.5 µg/µL RNA in 50 mM HEPES, pH 7.5, was heated at 80°C for 2 min, then 60°C for 2 min, and then $MgCl_2$ was added to a final concentration of 10 mM, followed by incubation on ice for a minimum of 30 min. Folded RNAs were then added to 8 mM SHAPE reagent, N-methylisatoic anhydride (NMIA) resuspended in dimethyl sulfoxide (DMSO), or just DMSO at a final concentration of 100 ng/µL for 45 minutes at 37°C. Reactions were quenched by ethanol precipitation and the RNAs were pelleted. The reverse transcription reaction was then performed on the resuspended RNA pellet using 6-carboxyfluorescein (6-FAM) labeled primers from Sigma-Aldrich and Superscript III reverse transcriptase. The reaction followed Invitrogen's suggested protocol. Lastly, 1 µL of reverse transcribed product was ethanol precipitated and submitted to the Ohio State Plant-Microbe Genomics Facility for capillary electrophoresis (CE) analysis. CE data was then analyzed using Microsoft Excel-based tools developed by our lab called RiboCAT and RiboDOG (Cantara et al.

2017). These tools simplify CE analysis and reduce manual intervention steps, allowing for a more precise output of overall SHAPE data (Cantara et al. 2017). After analysis, SHAPE reactivities are used as constraints to aid in the secondary structure folding of RNAs using RNAstructure (Mathews 2006).

RESULTS AND DISCUSSION

The lowest energy secondary structure of the HIV-1 400mer WT construct folds in accordance with previously solved 5'UTR structures from both our lab and the Weeks lab (Cantara et al. 2017, Wilkinson et al. 2008) (Figure 12A). The TAR and polyA form independent hairpins followed by the PBS stem loop and SL1-SL3 of Psi. Given the differences observed in salt titration assays with GagΔp6, we initially hypothesized that the secondary structure of the 400mer WT would be different from the secondary structures formed from the 356mer construct solved in (Cantara et al. 2017). Similarities in the two 5'UTR structures can be explained by the fact that SHAPE data is unable to probe the most 5' and 3' ends due to high intensity peaks corresponding to the full-length transcript and reverse transcriptase stalling during initiation of primer extension (Wilkinson et al. 2006). Thus, in the case with the 400mer WT, SHAPE reactivity constraints are lacking for the last 40 nt. To address this issue, future SHAPE studies will be performed on a longer gRNA construct allowing proper data collection of the region immediately downstream of Psi.

The SHAPE-derived secondary structure for the Vpr 400mer is shown in Figure 12B. Interestingly, while the TAR hairpin forms independently, the polyA hairpin base pairs with regions downstream of the major 5' splice site, which are unique to the Vpr mRNA transcript. The

PBS region and SL1 maintain the same structure as the 400mer WT gRNA. The exact function of the polyA hairpin in the 5'UTR is not entirely clear, but structural studies have found that destabilization of the polyA hairpin produces defects in both, RNA packaging and viral replication (Clever, Eckstein and Parslow 1999, Das et al. 1997, Smyth et al. 2018). Moreover, the secondary structure of the polyA hairpin rather than the primary sequence was found to be an important component for specific packaging (Rye-McCurdy et al. 2016, Clever et al. 1999). In the lowest energy SHAPE-predicted structure, the polyA hairpin is stabilized by 5 Watson-Crick base pairs, whereas in the WT 400mer construct, the polyA hairpin is stabilized by 16 base pairs. This drastic destabilization of the hairpin and the lack of an independent hairpin downstream of the TAR stem loop potentially explains how this mRNA transcript avoids being specifically packaged.

The SHAPE probed secondary structure of the Tat 400mer displays modest differences in TAR and polyA as compared to the 400mer WT RNA (Figure 12C). Although the hairpins are connected by an internal bulge, they maintain stable bases that stabilize their structure. In contrast, major changes are observed in the PBS domain. The PBS domain base pairs with 3' regions unique to the Tat 400mer, in contrast to the largely single-stranded structure of the PBS domain adopted by the 400mer WT. Strikingly, the exact 18-nt PBS sequence where the primer for reverse transcription, $tRNA^{Lys3}$, anneals is highly base paired and unreactive to the SHAPE reagent. The annealing of $tRNA^{Lys3}$ to the 18-nt PBS has been shown to be mediated by the NC domain of Gag (Feng et al. 1999, Jones et al. 2013). The process of tRNA annealing may be disturbed in the Tat 400mer construct, due to a stabilized structure of the PBS. These conformational differences may prevent tRNA annealing, which may be a pre-requisite for RNA packaging.

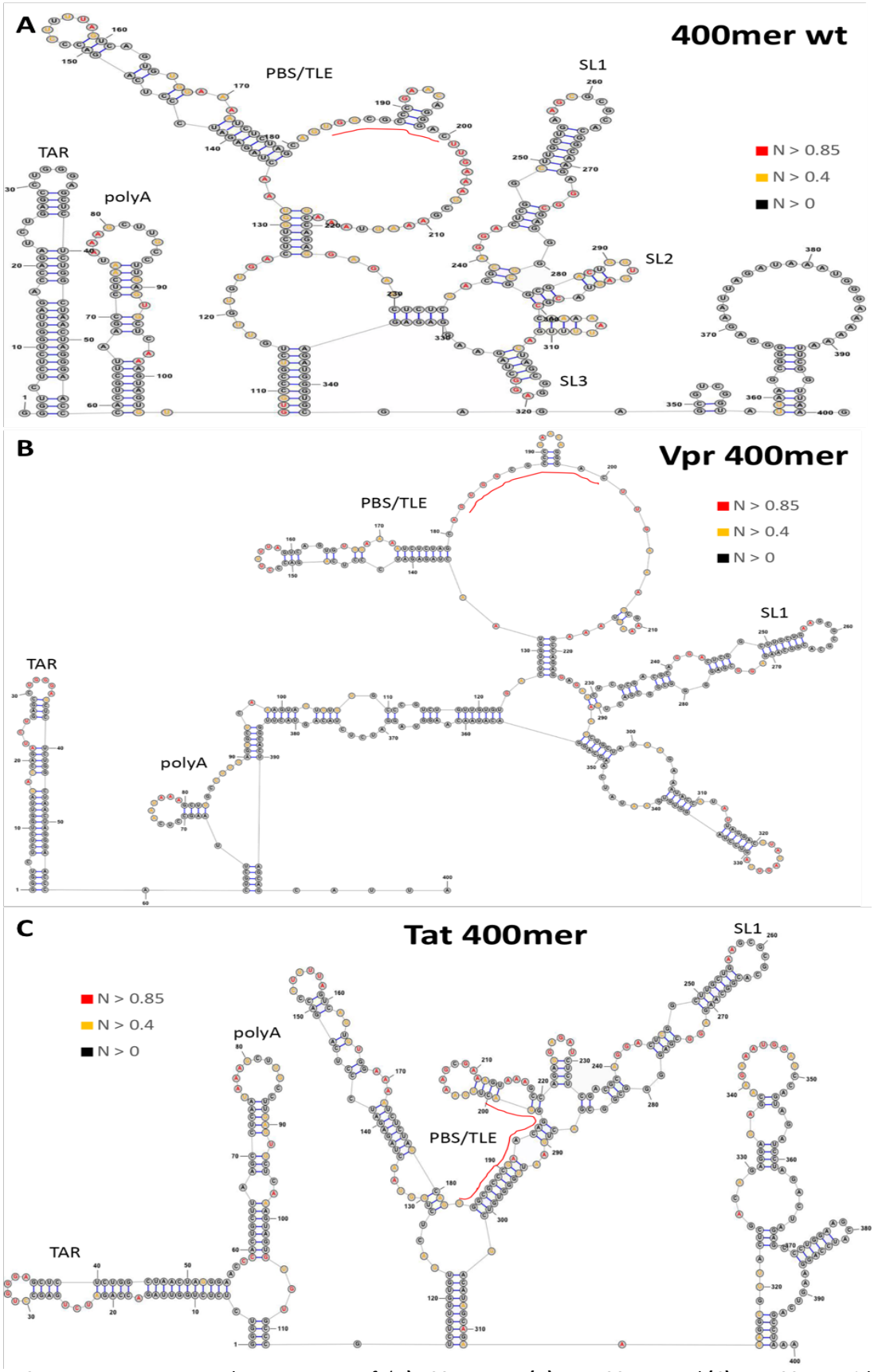


FIGURE 12. Lowest energy secondary structures of: (**A)** 400mer WT, **(B)** Vpr 400mer, and **(C)** Tat 400mer with SHAPE reactivities. Secondary structures were plotted via RNAstructure. The red line in the secondary structures indicates the 18-nt PBS.

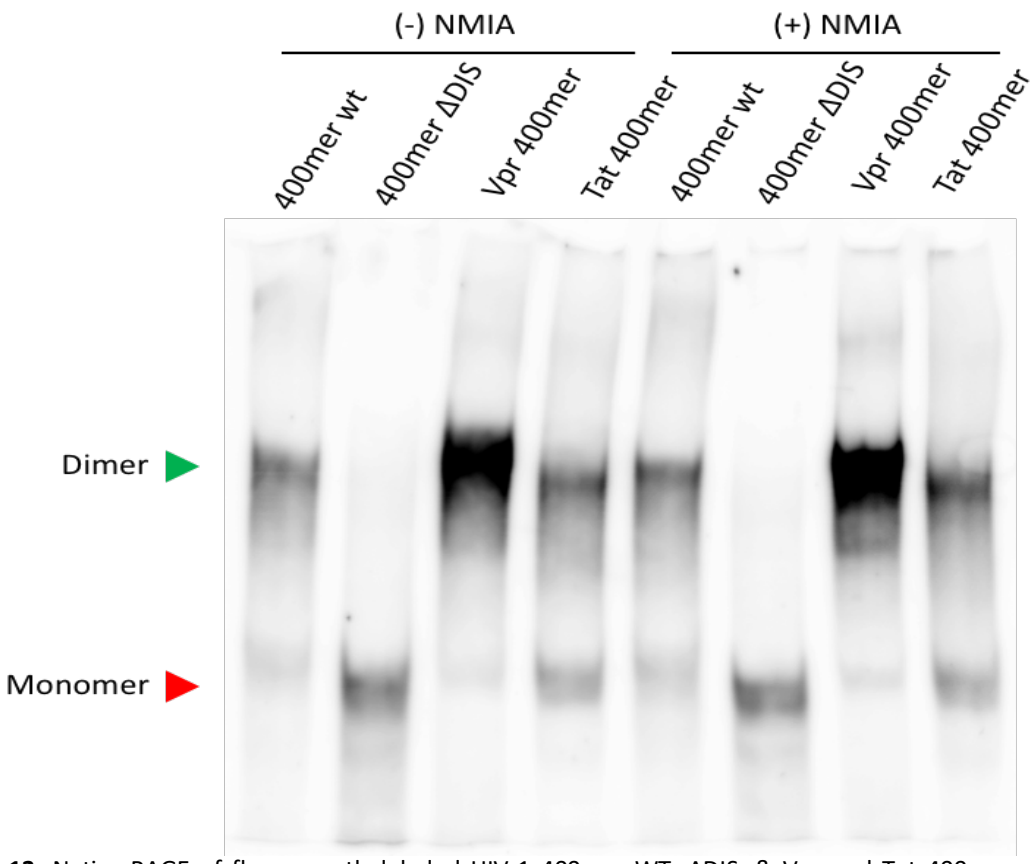


FIGURE 13. Native-PAGE of fluorescently-labeled HIV-1 400mer WT, ΔDIS, & Vpr and Tat 400mers after SHAPE reaction. RNAs were labeled as described in the main text and folded as follows: 0.5 µg/µL RNA in 50 mM HEPES, pH 7.5, was heated at 80°C for 2 min, then 60°C for 2 min, and then $MgCl_2$ was added to a final concentration of 10 mM, followed by incubation on ice for a minimum of 30 min. SHAPE reactions were then performed on the RNAs both with (+) and without (-) NMIA. After the 45-minute SHAPE incubation at 37°C, Folded RNAs were run on a 4.5% native-polyacrylamide gel (19:1 acrylamide: bis-acrylamide) also containing 1 mM $MgCl_2$ and TB (18 mM Tris, 9 mM borate, pH 8) at 4°C in TB running buffer. The final concentration HIV-1 RNA run on the gels shown was 83.25 ng/µL (617 nM) 2. Gels were visualized using the fluorescence mode on an AlphaImager EP (AlphaInnotech).

The overall global fold of the 400mer RNAs were assessed via native-PAGE to determine potential folding differences caused by the Vpr and Tat coding sequences downstream of the major 5' splice site. The RNAs were run on a native-PAGE after the SHAPE reaction to properly assess their conformation in our SHAPE study (Figure 13). The 400mer WT forms two distinct bands with the major band corresponding to the dimer RNA and the minority band corresponding to the monomer. Identities of the monomer and dimer bands can be distinguished by comparison to the dimer-incompetent, 400mer ΔDIS monomer control. The Vpr 400mer forms a dominant

dimer band with very little monomer product. On the other hand, the Tat 400mer forms equal monomer and dimer bands. In all of these constructs, the DIS loop in SL1 is exposed and able to dimerize but other interactions may be affecting the ability of the 400mer WT and Tat 400mer to fully dimerize. Future studies will address this discrepancy by performing SHAPE reactions on dimer-incompetent (ΔDIS) 400mer WT, Vpr 400mer, and Tat 400mer.

Overall, our SHAPE studies of the 400-nt vsRNAs, Vpr and Tat, revealed structural differences compared to a 400-nt gRNA, 400mer WT. We explored the implication of these disrupted structures and their potential effect on specific packaging and concluded that each vsRNA construct that we studied has a potential mechanism by which it avoids specific packaging into virions. In the Vpr 400mer, we found that the polyA hairpin was destabilized, which has been shown to reduce gRNA packaging and viral replication. In the Tat 400mer, we found that the 18-nt PBS was sequestered and unreactive, which may prevent $tRNA^{Lys3}$ annealing. The exact step of the HIV-1 lifecycle when tRNA anneals is unknown but is has been proposed that annealing may be important for dimerization and packaging (Seif, Niu and Kleiman 2013).

ACKNOWLEDGEMENTS

I want to thank Erik Olson for designing the Vpr and Tat 400mer plasmid constructs. Also, I would to thank Nathasha Hewagama for her contributions in preparing these RNA constructs and performing the SHAPE reactions.

CONCLUSIONS AND FUTURE DIRECTIONS

In this thesis, the mechanisms by which Gag selects its genomic RNA for packaging were examined. Specific guanosine residues in Psi, larger flanking regulatory elements within the 5'UTR, and spliced sequences ligated downstream of the major 5' splice site were all probed. Guanosine residues in the internal bulge of SL1 and the single-stranded region downstream of SL3 diminished Gag binding specificity, consistent with other studies. Furthermore, guanosine residues in SL2 and SL3 were not significant at reducing overall Gag specificity. The data reported here provide weak evidence in support of previously proposed upstream negative regulatory element, but sequences downstream of Psi are able to increase GagΔp6 binding specificity independent of the upstream element and its sequence. Since downstream sequences were sufficient to increase Gag specificity, how segments of vsRNAs that encode for Vpr and Tat avoid specific packaging was examined. SHAPE probing allowed us to solve the secondary structures of our vsRNA constructs, revealing potential anti-packaging mechanisms that rely on restructuring of important regulatory elements. Overall, these studies have furthered our insight into the complexity of genomic RNA packaging.

In our future studies, we plan to continue exploring the Vpr and Tat 400mer constructs by performing salt titration assays with GagΔp6, comparing their behavior to the 400mer WT. This will allow us to determine their overall Gag binding specificity, as well as the conformation in which Gag is binding to these RNAs. This will also directly address our hypothesis on the polyA hairpin destabilization in the Vpr 400mer construct and whether that impacts specific recognition by Gag. To address our hypothesis about the sequestered PBS in the Tat 400mer construct, we plan on performing NC-mediated $tRNA^{Lys3}$ annealing assays to determine differences in the ability

of NC to anneal $tRNA^{Lys3}$ to this construct and the 400mer WT.

Lastly, we plan on studying the effect of *trans*-acting elements on the 5'UTR and how they affect specific recognition by Gag. We are particularly interested in how $tRNA^{Lys3}$ modulates the structure of the 5'UTR upon annealing. Current literature suggests that annealing of $tRNA^{Lys3}$ to the PBS promotes dimerization of the RNA, which is then favored for packaging (Seif et al. 2013). Preliminary heat annealing assays suggests that dimerization is favored, but only slightly. To explore this further, we plan to perform NC-mediated annealing to properly assess the promotion of a specific conformational state. Furthermore, we plan to perform salt titration assays to determine Gag specificity.

REFERENCES

Abbink, T. E. & B. Berkhout (2003) A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. *J Biol Chem,* 278**,** 11601-11.

Abd El-Wahab, E. W., R. P. Smyth, E. Mailler, S. Bernacchi, V. Vivet-Boudou, M. Hijnen, F. Jossinet, J. Mak, J. C. Paillart & R. Marquet (2014) Specific recognition of the HIV-1 genomic RNA by the Gag precursor. *Nat Commun,* 5**,** 4304.

Adachi, A., H. E. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson & M. A. Martin (1986) Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J Virol,* 59**,** 284-91.

Aldovini, A. & R. A. Young (1990) Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. *J Virol,* 64**,** 1920-6.

Bannwarth, S. & A. Gatignol (2005) HIV-1 TAR RNA: the target of molecular interactions between the virus and its host. *Curr HIV Res,* 3**,** 61-71.

Berkowitz, R., J. Fisher & S. P. Goff (1996) RNA packaging. *Curr Top Microbiol Immunol,* 214**,** 177-218.

Bernacchi, S., E. W. Abd El-Wahab, N. Dubois, M. Hijnen, R. P. Smyth, J. Mak, R. Marquet & J. C. Paillart (2017) HIV-1 Pr55. *RNA Biol,* 14**,** 90-103.

Bryksin, A. V. & I. Matsumura (2010) Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques,* 48**,** 463-5.

Cantara, W. A., J. Hatterschide, W. Wu & K. Musier-Forsyth (2017) RiboCAT: a new capillary electrophoresis data analysis tool for nucleic acid probing. *RNA,* 23**,** 240-249.

Chen, J., O. Nikolaitchik, J. Singh, A. Wright, C. E. Bencsics, J. M. Coffin, N. Ni, S. Lockett, V. K. Pathak & W. S. Hu (2009) High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *Proc Natl Acad Sci U S A,* 106**,** 13535-40.

Chiu, J., P. E. March, R. Lee & D. Tillett (2004) Site-directed, Ligase-Independent Mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h. *Nucleic Acids Res,* 32**,** e174.

Chukkapalli, V. & A. Ono (2011) Molecular determinants that regulate plasma membrane association of HIV-1 Gag. *J Mol Biol,* 410**,** 512-24.

Clever, J. L., D. A. Eckstein & T. G. Parslow (1999) Genetic dissociation of the encapsidation and reverse transcription functions in the 5' R region of human immunodeficiency virus type 1. *J Virol,* 73**,** 101-9.

Cristofari, G. & J. L. Darlix (2002) The ubiquitous nature of RNA chaperone proteins. *Prog Nucleic Acid Res Mol Biol,* 72**,** 223-68.

D'Souza, V. & M. F. Summers (2005) How retroviruses select their genomes. *Nat Rev Microbiol,* 3**,** 643-55.

Das, A. T., B. Klaver, B. I. Klasens, J. L. van Wamel & B. Berkhout (1997) A conserved hairpin motif in the R-U5 region of the human immunodeficiency virus type 1 RNA genome is essential for replication. *J Virol,* 71**,** 2346-56.

Datta, S. A., F. Heinrich, S. Raghunandan, S. Krueger, J. E. Curtis, A. Rein & H. Nanda (2011) HIV-1 Gag extension: conformational changes require simultaneous interaction with membrane and nucleic acid. *J Mol Biol,* 406**,** 205-14.

Datta, S. A. & A. Rein (2009) Preparation of recombinant HIV-1 gag protein and assembly of virus-like particles in vitro. *Methods Mol Biol,* 485**,** 197-208.

Feng, Y. X., S. Campbell, D. Harvin, B. Ehresmann, C. Ehresmann & A. Rein (1999) The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. *J Virol,* 73**,** 4251-6.

Goldschmidt, V., M. Rigourd, C. Ehresmann, S. F. Le Grice, B. Ehresmann & R. Marquet (2002) Direct and indirect contributions of RNA secondary structure elements to the initiation of HIV-1 reverse transcription. *J Biol Chem,* 277**,** 43233-42.

Heng, X., S. Kharytonchyk, E. L. Garcia, K. Lu, S. S. Divakaruni, C. LaCotti, K. Edme, A. Telesnitsky & M. F. Summers (2012) Identification of a minimal region of the HIV-1 5'-leader required for RNA dimerization, NC binding, and packaging. *J Mol Biol,* 417**,** 224-39.

Houzet, L., J. C. Paillart, F. Smagulova, S. Maurel, Z. Morichaud, R. Marquet & M. Mougel (2007) HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. *Nucleic Acids Res,* 35**,** 2695-704.

Jones, C. P., W. A. Cantara, E. D. Olson & K. Musier-Forsyth (2014) Small-angle X-ray scattering-derived structure of the HIV-1 5' UTR reveals 3D tRNA mimicry. *Proc Natl Acad Sci U S A,* 111**,** 3395-400.

Jones, C. P., S. A. Datta, A. Rein, I. Rouzina & K. Musier-Forsyth (2011) Matrix domain modulates HIV-1 Gag's nucleic acid chaperone activity via inositol phosphate binding. *J Virol,* 85**,** 1594-603.

Jones, C. P., J. Saadatmand, L. Kleiman & K. Musier-Forsyth (2013) Molecular mimicry of human tRNALys anti-codon domain by HIV-1 RNA genome facilitates tRNA primer annealing. *RNA,* 19**,** 219-29.

Keane, S. C., X. Heng, K. Lu, S. Kharytonchyk, V. Ramakrishnan, G. Carter, S. Barton, A. Hosic, A. Florwick, J. Santos, N. C. Bolden, S. McCowin, D. A. Case, B. A. Johnson, M. Salemi, A. Telesnitsky & M. F. Summers (2015) RNA structure. Structure of the HIV-1 RNA packaging signal. *Science,* 348**,** 917-21.

Kutluay, S. B., T. Zang, D. Blanco-Melo, C. Powell, D. Jannain, M. Errando & P. D. Bieniasz (2014) Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. *Cell,* 159**,** 1096-1109.

Kuzembayeva, M., K. Dilley, L. Sardo & W. S. Hu (2014) Life of psi: how full-length HIV-1 RNAs become packaged genomes in the viral particles. *Virology,* 454-455**,** 362-70.

Levin, J. G., J. Guo, I. Rouzina & K. Musier-Forsyth (2005) Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. *Prog Nucleic Acid Res Mol Biol,* 80**,** 217-86.

Lu, K., X. Heng & M. F. Summers (2011) Structural determinants and mechanism of HIV-1 genome packaging. *J Mol Biol,* 410**,** 609-33.

Mak, J. & L. Kleiman (1997) Primer tRNAs for reverse transcription. *J Virol,* 71**,** 8087-95.

Martin-Serrano, J. & S. J. Neil (2011) Host factors involved in retroviral budding and release. *Nat Rev Microbiol,* 9**,** 519-31.

Mathews, D. H. (2006) RNA secondary structure analysis using RNAstructure. *Curr Protoc Bioinformatics,* Chapter 12**,** Unit 12.6.

McBride, M. S. & A. T. Panganiban (1996) The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. *J Virol,* 70**,** 2963-73.

--- (1997) Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo. *J Virol,* 71**,** 2050-8.

Meng, B. & A. M. Lever (2013) Wrapping up the bad news: HIV assembly and release. *Retrovirology,* 10**,** 5.

Milligan, J. F., D. R. Groebe, G. W. Witherell & O. C. Uhlenbeck (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res,* 15**,** 8783-98.

Moore, M. D. & W. S. Hu (2009) HIV-1 RNA dimerization: It takes two to tango. *AIDS Rev,* 11**,** 91-102.

Munro, J. B., A. Nath, M. Färber, S. A. Datta, A. Rein, E. Rhoades & W. Mothes (2014) A conformational transition observed in single HIV-1 Gag molecules during in vitro assembly of virus-like particles. *J Virol,* 88**,** 3577-85.

Musumeci, D., C. Riccardi & D. Montesarchio (2015) G-Quadruplex Forming Oligonucleotides as Anti-HIV Agents. *Molecules,* 20**,** 17511-32.

Nikolaitchik, O. A., K. A. Dilley, W. Fu, R. J. Gorelick, S. H. Tai, F. Soheilian, R. G. Ptak, K. Nagashima, V. K. Pathak & W. S. Hu (2013) Dimeric RNA recognition regulates HIV-1 genome packaging. *PLoS Pathog,* 9**,** e1003249.

O'Carroll, I. P., R. M. Crist, J. Mirro, D. Harvin, F. Soheilian, A. Kamata, K. Nagashima & A. Rein (2012) Functional redundancy in HIV-1 viral particle assembly. *J Virol,* 86**,** 12991-6.

O'Carroll, I. P., F. Soheilian, A. Kamata, K. Nagashima & A. Rein (2013) Elements in HIV-1 Gag contributing to virus particle assembly. *Virus Res,* 171**,** 341-5.

Pagano, J. M., B. M. Farley, L. M. McCoig & S. P. Ryder (2007) Molecular basis of RNA recognition by the embryonic polarity determinant MEX-5. *J Biol Chem,* 282**,** 8883-94.

Rein, A., S. A. Datta, C. P. Jones & K. Musier-Forsyth (2011) Diverse interactions of retroviral Gag proteins with RNAs. *Trends Biochem Sci,* 36**,** 373-80.

Rein, A., L. E. Henderson & J. G. Levin (1998) Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. *Trends Biochem Sci,* 23**,** 297-301.

Rulli, S. J., Jr., C. S. Hibbert, J. Mirro, T. Pederson, S. Biswal & A. Rein (2007) Selective and nonselective packaging of cellular RNAs in retrovirus particles. *J Virol,* 81**,** 6623-31.

Rye-McCurdy, T., E. D. Olson, S. Liu, C. Binkley, J. P. Reyes, B. R. Thompson, J. M. Flanagan, L. J. Parent & K. Musier-Forsyth (2016) Functional Equivalence of Retroviral MA Domains in Facilitating Psi RNA Binding Specificity by Gag. *Viruses,* 8.

Rye-McCurdy, T., I. Rouzina & K. Musier-Forsyth (2015) Fluorescence anisotropy-based salt-titration approach to characterize protein-nucleic acid interactions. *Methods Mol Biol,* 1259**,** 385-402.

Schur, F. K., W. J. Hagen, M. Rumlová, T. Ruml, B. Müller, H. G. Kräusslich & J. A. Briggs (2015) Structure of the immature HIV-1 capsid in intact virus particles at 8.8 Å resolution. *Nature,* 517**,** 505-8.

Seif, E., M. Niu & L. Kleiman (2013) Annealing to sequences within the primer binding site loop promotes an HIV-1 RNA conformation favoring RNA dimerization and packaging. *RNA,* 19**,** 1384-93.

Smyth, R. P., L. Despons, G. Huili, S. Bernacchi, M. Hijnen, J. Mak, F. Jossinet, L. Weixi, J. C. Paillart, M. von Kleist & R. Marquet (2015) Mutational interference mapping experiment (MIME) for studying RNA structure and function. *Nat Methods,* 12**,** 866-72.

Smyth, R. P., M. R. Smith, A. C. Jousset, L. Despons, G. Laumond, T. Decoville, P. Cattenoz, C. Moog, F. Jossinet, M. Mougel, J. C. Paillart, M. von Kleist & R. Marquet (2018) In cell mutational interference mapping experiment (in cell MIME) identifies the 5' polyadenylation signal as a dual regulator of HIV-1 genomic RNA production and packaging. *Nucleic Acids Res*.

Stoltzfus, C. M. (2009) Chapter 1. Regulation of HIV-1 alternative RNA splicing and its role in virus replication. *Adv Virus Res,* 74**,** 1-40.

Webb, J. A., C. P. Jones, L. J. Parent, I. Rouzina & K. Musier-Forsyth (2013) Distinct binding interactions of HIV-1 Gag to Psi and non-Psi RNAs: implications for viral genomic RNA packaging. *RNA,* 19**,** 1078-88.

Wilkinson, K. A., R. J. Gorelick, S. M. Vasa, N. Guex, A. Rein, D. H. Mathews, M. C. Giddings & K. M. Weeks (2008) High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol,* 6**,** e96.

Wilkinson, K. A., E. J. Merino & K. M. Weeks (2006) Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat Protoc,* 1**,** 1610-6.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res,* 31**,** 3406-15.